

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :  C12P 21/06, C12N 15/00	A1	(11) International Publication Number: WO 91/06666  (43) International Publication Date: 16 May 1991 (16.05.91)
--	----	---

(21) International Application Number: PCT/US90/06425  (22) International Filing Date: 6 November 1990 (06.11.90)  (30) Priority data: 432,069 6 November 1989 (06.11.89) US  (71) Applicant: CELL GENESYS, INC. [US/US]; 344 Ladesside Drive, Foster City, CA 94404 (US).  (72) Inventor: SKOULTCHI, Arthur, I. ; 71 N. Chatsworth Avenue, Larchmont, NY 10538 (US).  (74) Agents: ROWLAND, Bertram, I. et al.; Cooley Godward Castro Huddeson & Tatum, Five Palo Alto Square, 4th Floor, Palo Alto, CA 94306 (US).	(81) Designated State: KR.  Published <i>With international search report.</i>
---	---

(54) Title: PRODUCTION OF PROTEINS USING HOMOLOGOUS RECOMBINATION

(57) Abstract

Methods and compositions are provided for expression of mammalian genes in culture. An amplifiable gene is introduced by homologous recombination in juxtaposition to a target gene, the resulting combination of amplifiable gene and target gene transferred to a convenient host and the target gene amplified by means of the amplifiable gene. The resulting expression host may then be grown in culture with enhanced expression of the target gene.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MR	Mauritania
BE	Belgium	GA	Gabon	MW	Malawi
BF	Burkina Faso	GB	United Kingdom	NL	Netherlands
BG	Bulgaria	GR	Greece	NO	Norway
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	SD	Sudan
CF	Central African Republic	KP	Democratic People's Republic of Korea	SE	Sweden
CG	Congo	KR	Republic of Korea	SN	Senegal
CH	Switzerland	LJ	Liechtenstein	SU	Soviet Union
CI	Côte d'Ivoire	LK	Sri Lanka	TD	Chad
CM	Cameroon	LU	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	US	United States of America
DK	Denmark				

## PRODUCTION OF PROTEINS

## 5 USING HOMOLOGOUS RECOMBINATION

INTRODUCTION

10

Technical Field

The field of this invention is the expression of mammalian proteins.

15

Background

20

The discoveries of restriction enzymes, cloning, sequencing, reverse transcriptase, and monoclonal antibodies has resulted in extraordinary capabilities in isolating, identifying, and manipulating nucleic acid sequences. As a result of these capabilities, numerous genes and their transcriptional control elements have been identified and manipulated. The genes have been used for producing large amounts of a desired protein in heterologous hosts (bacterial and eukaryotic host cell systems).

25

In many cases, the process of obtaining coding sequences and eliciting their expression has been a long and arduous one. The identification of the coding sequence, either cDNA or genomic DNA, has frequently involved the construction of libraries, identification of fragments of the open reading frame, examining the flanking sequences, and the like. In mammalian genes where introns are frequently encountered, in many instances, the coding region has been only small fraction of the total nucleic acid associated with the gene. In other cases, pseudogenes or multi-membered gene families have obscured the ability to isolate a particular gene of interest.

30

35

Nevertheless, as techniques have improved, there has

been a continuous parade of successful identifications and isolation of genes of interest.

In many situations one is primarily interested in a source of the protein product. The 5 cell type in the body which produces the protein is frequently an inadequate source, since the protein may be produced in low amounts, the protein may only be produced in a differentiated host cell which is only difficultly grown in culture, or the host cell, 10 particularly a human cell, is not economic or efficient in a culture process for production of the product. There is, therefore, significant interest in developing alternative techniques for producing proteins of interest in culture with cells which 15 provide for economic and efficient production of the desired protein and, when possible, appropriate processing of the protein product.

#### Relevant Literature

20 Mansour *et al.*, Nature, 336:348-352 (1988), describe a general strategy for targeting mutations to non-selectable genes. Weidle *et al.*, Gene, 66:193-203, (1988), describe amplification of tissue-type plasminogen activator with a DHFR gene and loss of 25 amplification in the absence of selective pressure. Murnane and Yezzi, Somatic Cell and Molecular Genetics, 14:273-286, (1988), describe transformation of a human cell line with an integrated selectable gene marker lacking a transcriptional promoter, with 30 tandem duplication and amplification of the gene marker. Thomas and Capecchi, Cell, 51:503-512, (1987), describe site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Song *et al.*, Proc. Natl. Acad. Sci. USA, 84:6820-6824, 35 (1987), describe homologous recombination in human cells by a two staged integration. Liskay *et al.*, "Homologous Recombination Between Repeated Chromosomal Sequences in Mouse Cells," Cold Spring Harbor, Symp.

Quant. Biol. 49:13-189, (1984), describe integration of two different mutations of the same gene and homologous recombination between the mutant genes. Rubnitz and Subramani, Mol. and Cell. Biol. 4:2253-2258, (1984), describe the minimum amount of homology required for homologous recombination in mammalian cells. Kim and Smithies, Nucl. Acids. Res. 16:8887-8903, (1988), describe an assay for homologous recombination using the polymerase chain reaction.

10

#### SUMMARY OF THE INVENTION

Expression of mammalian proteins of interest is achieved by employing homologous recombination for integration of an amplifiable gene and other regulatory sequences in proximity to a gene of interest without interruption of the production of a proper transcript. The region comprising the amplifiable gene and the gene of interest may be amplified, the genome fragmented and directly or indirectly transferred to an expression host for expression of the target protein. If not previously amplified, the target region is then amplified, and the cell population screened for cells producing the target protein. Cells which produce the target protein at high and stable levels are expanded and used for expression of the target protein.

#### BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a diagrammatic illustration of the plasmid pCG.1 showing the sequence of the modified polylinker;

FIG. 2 is a diagrammatic illustration of the construction of the plasmid pCG.HR1;

FIG. 3 is a diagrammatic illustration of the result of targeting the EPO locus by homologous recombination with the DNA from pCG.HR1 cut with NotI;

FIG. 4 is a diagrammatic illustration of the PCR amplification fragment produced from cells in which a homologous recombination event has occurred.

5

#### DESCRIPTION OF SPECIFIC EMBODIMENTS

Methods and compositions are provided for production of mammalian proteins of interest in culture. The method employs homologous recombination in a host cell for integrating an amplifiable gene in the vicinity of a target gene, which target gene encodes the protein of interest. The region comprising both the amplifiable gene and target gene will be referred to as the amplifiable region. The resulting transformed primary cells may now be subjected to conditions which select for amplification, or the amplification may be performed subsequently. "Transform" includes transform, transfect, transduce, conjugation, fusion, electroporation or any other technique for introducing DNA into a viable cell. The chromosomes or DNA of the transformed cells are then used to transfer the amplifiable region into the genome of secondary expression host cells, where the target region, if not previously amplified sufficiently or at all, is further amplified. The resulting cell lines are screened for production of the target protein and secondary cell lines selected for desired levels of production, which cells may be expanded and used for production of the desired protein in culture.

30

The primary cell may be any mammalian cell of interest, particularly mammalian cells which do not grow readily in culture, more particularly primate cells, especially human cells, where the human cells may be normal cells, including embryonic or neoplastic cells, particularly normal cells. Various cell types may be employed as the primary cells, including fibroblasts, particularly diploid skin fibroblasts, lymphocytes, epithelial cells, neurons, endothelial

cells, or other somatic cells, or germ cells. Of particular interest are skin fibroblasts, which can be readily propagated to provide for large numbers of normal cells, embryonic kidney cells, and the like.

5 These cells may or may not be expressing the gene of interest. In those instances where the target gene is inducible or only expressed in certain differentiated cells, one may select cells in which the target gene is expressed, which may require immortalized cells  
10 capable of growth in culture.

A number of amplifiable genes exist, where by appropriate use of a selection agent, a gene integrated in the genome will be amplified with adjacent flanking DNA. Amplifiable genes include  
15 dihydrofolate reductase, metallothionein-I and -II, preferably primate metallothionein genes, adenosine deaminase, ornithine decarboxylase, etc. The amplifiable gene will have transcriptional signals which are functional in the secondary or expression  
20 host and desirably be functional in the primary host, particularly where amplification is employed in the primary host or the amplifiable gene is used as a marker.

The target genes may be any gene of  
25 interest, there already having been a large number of proteins of interest identified and isolated with continual additions to the list. Proteins of interest include cytokines, such as interleukins 1-10; growth factors such as EGF, FGF, PDGF, and TGF;  
30 somatotropins; growth hormones; colony stimulating factors, such as G-, M-, and GM-CSF; erythropoietin; plasminogen activators, such as tissue and urine; enzymes, such as superoxide dismutase; interferons; T-cell receptors; surface membrane proteins; insulin;  
35 lipoproteins;  $\alpha_1$ -antitrypsin; CD proteins, such as CD3, 4, 8, 19; clotting factors, e.g., Factor VIIIic and von Will brands factor; anticoagulation factors, such as Protein C; atrial natriuretic factor, tumor necrosis

factor; transport proteins; homing receptors;  
addressins; regulatory proteins; etc.

For homologous recombination, constructs will be prepared where the amplifiable gene will be flanked on one or both sides with DNA homologous with the DNA of the target region. The homologous DNA will generally be within 100 kb, usually 50 kb, preferably about 25 kb, of the transcribed region of the target gene, more preferably within 2 kb of the target gene.

By gene is intended the coding region and those sequences required for transcription of a mature mRNA. The homologous DNA may include the 5'-upstream region comprising any enhancer sequences, transcriptional initiation sequences, the region 5' of these sequences, or the like. The homologous region may include a portion of the coding region, where the coding region may be comprised only of an open reading frame or combination of exons and introns. The homologous region may comprise all or a portion of an intron, where all or a portion of one or more exons may also be present. Alternatively, the homologous region may comprise the 3'-region, so as to comprise all or a portion of the transcription termination region, or the region 3' of this region. The homologous regions may extend over all or a portion of the target gene or be outside the target gene comprising all or a portion of the transcriptional regulatory regions and/or the structural gene. For the most part, the homologous sequence will be joined to the amplifiable gene, proximally or distally. Usually a sequence other than the wild-type sequence normally associated with the target gene will be used to separate the homologous sequence from the amplifiable gene on at least one side of the amplifiable gene. Some portion of the sequence may be the 5' or 3' sequence associated with the amplifiable gene, as a result of the manipulations associated with the amplifiable gene.

The homologous regions flanking the amplifiable gene need not be identical to the target region, where in vitro mutagenesis is desired. For example, one may wish to change the transcriptional initiation region for the target gene, so that a portion of the homologous region might comprise nucleotides different from the wild-type 5' region of the target gene. Alternatively, one could provide for insertion of a transcriptional initiation region different from the wild-type initiation region between the wild-type initiation region and the structural gene. Similarly, one might wish to introduce various mutations into the structural gene, so that the homologous region would comprise mismatches, resulting in a change in the encoded protein. For example, a signal leader sequence would be introduced in proper reading frame with the target gene to provide for secretion of the target protein expression product. Alternatively, one might change the 3' region, e.g., untranslated region, polyadenylation site, etc. of the target gene. Therefore, by homologous recombination, one can provide for maintaining the integrity of the target gene, so as to express the wild-type protein under the transcriptional regulation of the wild-type promoter or one may provide for a change in transcriptional regulation, processing or sequence of the target gene. In some instances, one may wish to introduce an enhancer in relation to the transcriptional initiation region, which can be provided by, for example, integration of the amplifiable gene associated with the enhancer in a region upstream from the transcriptional initiation regulatory region or in an intron or even downstream from the target gene.

In order to prepare the subject constructs, it will be necessary to know the sequence which is targeted for homologous recombination. While it is

reported that a sequence of 14 bases complementary to a sequence in a genome may provide for homologous recombination, normally the individual flanking sequences will be at least about 150 bp, and may be 5 12 kb or more, usually not more than about 8 kb. The size of the flanking regions will be determined by the size of the known sequence, the number of sequences in the genome which may have homology to the site for integration, whether mutagenesis is involved and the 10 extent of separation of the regions for mutagenesis, the particular site for integration, or the like.

The integrating constructs may be prepared in accordance with conventional ways, where sequences may be synthesized, isolated from natural sources, 15 manipulated, cloned, ligated, subjected to in vitro mutagenesis, primer repair, or the like. At various stages, the joined sequences may be cloned, and analyzed by restriction analysis, sequencing, or the like. Usually the construct will be carried on a 20 cloning vector comprising a replication system functional in a prokaryotic host, e.g., E. coli, and a marker for selection, e.g., biocide resistance, complementation to an auxotrophic host, etc. Other functional sequences may also be present, such as 25 polylinkers, for ease of introduction and excision of the construct or portions thereof, or the like. A large number of cloning vectors are available such as pBR322, the pUC series, etc.

Once the construct is prepared, it may then 30 be used for homologous recombination in the primary cell target. Various techniques may be employed for integrating the construct into the genome of the primary cell without being joined to a replication system functional in the primary host. See for 35 example, U.S. Patent No. 4,319,216, as well as the references cited in the Relevant Literature section. Alternatively, the construct may be inserted into an

appropriate vector, usually having a viral replication system, such as SV40, bovine papilloma virus, adenovirus, or the like. The linear DNA sequence vector may also have a selectable marker for identifying transfected cells. Selectable markers include the neo gene, allowing for selection with G418, the herpes tk gene for selection with HAT medium, gpt gene with mycophenolic acid, complementation of an auxotrophic host, etc.

The vector may or may not be capable of stable maintenance in the host. Where the vector is capable of stable maintenance, the cells will be screened for homologous integration of the vector into the genome of the host, where various techniques for curing the cells may be employed. Where the vector is not capable of stable maintenance, for example, where a temperature sensitive replication system is employed, one may change the temperature from the permissive temperature to the non-permissive temperature, so that the cells may be cured of the vector. In this case, only those cells having integration of the construct comprising the amplifiable gene and, when present, the selectable marker, will be able to survive selection.

Where a selectable marker is present, one may select for the presence of the construct by means of the selectable marker. Where the selectable marker is not present, one may select for the presence of the construct by the amplifiable gene. For the neo gene or the herpes tk gene, one could employ a medium for growth of the transformants of about 0.1-1 g/ml of G418 or HAT medium respectively. Where DHFR is the amplifiable gene, the selective medium may include from about 0.01-0.25  $\mu$ M of methotrexate.

In carrying out the homologous recombination, the DNA will be introduced into the primary cells. Techniques which may be used include calcium phosphate/DNA co-precipitates, microinjection

of DNA into the nucleus, electroporation, bacterial protoplast fusion with intact cells, transfection, polycations, e.g., polybrene, polyornithine, etc., or the like. The DNA may be single or double stranded 5 DNA, linear or circular. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology (1989), Keown et al., Methods and Enzymology (1990) Vol. 185, pp. 527-537 and Mansour et al., Nature, 336:348-352, (1988).

10 Upstream and/or downstream from the target region construct may be a gene which provides for identification of whether a double crossover has occurred. For this purpose, the herpes simplex virus thymidine kinase gene may be employed since the 15 presence of the thymidine kinase gene may be detected by the use of nucleoside analogs, such as acyclovir or gancyclovir, for their cytotoxic effects on cells that contain a functional HSV-tk gene. The absence of sensitivity to these nucleoside analogs indicates the 20 absence of the thymidine kinase and, therefore, where homologous recombination has occurred, that a double crossover event has also occurred.

The presence of the marker gene as evidenced 25 by resistance to a biocide or growth in a medium which selects for the presence of the marker gene, establishes the presence and integration of the target construct into the host genome. No further selection need be made at this time, since the selection will be made in the secondary expression host, where 30 expression of the amplified target gene may be detected. If one wishes, one can determine whether homologous recombination has occurred by employing PCR and sequencing the resulting amplified DNA sequences. If desired, amplification may be performed at this 35 time by stressing the primary cells with the appropriate amplifying reagent, so that multi-copies of the target gene are obtained. Alternatively,

amplification may await transfer to th secondary cell expression host.

High molecular weight DNA, greater than about 20kb, preferably greater than about 50kb DNA or preferably metaphase chromosomes are prepared from the primary recipient cell strain having the appropriate integration of the amplification vector. Preparation and isolation techniques are described by Nelson and Housman, In Gene Transfer (ed. R. Kucherlapati) Plenum Press, 1986. The DNA may then be introduced in the same manner as described above into the secondary host expression cells, using the same or different techniques than employed for the primary cells. Various mammalian expression hosts are available and may be employed. These hosts include CHO cells, monkey kidney cells, C127 mouse fibroblasts, 3T3 mouse cells, Vero cells, etc. Desirably the hosts will have a negative background for the amplifiable gene or a gene which is substantially less responsive to the amplifying agent.

The transformed cells are grown in selective medium containing about 0.01-0.5  $\mu$ M methotrexate and, where another marker is present, e.g., the neo gene, the medium may contain from about 0.1-1 mg/ml G418. The resistant colonies are isolated and may then be analyzed for the presence of the construct in juxtaposition to the target gene. This may be as a result of detection of expression of the target gene product, where there will normally be a negative background for the target gene product, use of PCR, Southern hybridization, or the like.

The cells containing the construct are then expanded and subjected to selection and amplification with media containing progressively higher concentrations of the amplifying reagent, for example, 0.5-200  $\mu$ M of m thotrxate for the DHFR gene, and may b analyzed at each selection step for production of th target product. Expansion will

includ at least duplication and may result in at least 5 copies, preferably 10 copies or more in a tandem relationship. Thus protein production will be increased at least 1.5 fold from expression from a 5 single copy, usually at least 3 fold, preferably at least 5 fold.

The various clones may then be screened for optimum stable production of the target product and these clones may then be expanded and used 10 commercially for production in culture. In this manner, high yields of a product may be obtained, without the necessity of isolating the message and doing the various manipulations associated with genetic engineering or isolating the genomic gene, 15 where very large genes can be a major research and development effort.

The following examples are offered by way of illustration and not by way of limitation.

20 EXPERIMENTAL

Cells

Normal human diploid skin fibroblasts, ("primary recipient") are propagated in EEMEM medium 25 supplemented with 20% fetal calf serum. Dihydrofolate reductase (DHFR) deficient Chinese hamster ovary (CHO) DUKX-B11 cells (Urlaub and Chasin, Proc. Natl. Acad. Sci. USA 77:4216-4220 (1980)) ("secondary recipient") are propagated in alpha-medium supplemented with 10% 30 dialyzed fetal bovine serum.

DNA Vector

The amplification vector is constructed from pUC19 (Yanisch-Perron et al., Gene 33:103-119 35 (1985)). A 1.8 kb HaeII fragment containing a hygromycin B phosphotransferase gene (hph) driven by the herpes simplex virus thymidine kinase (HSV tk) promoter is isolated from pHg (Sugden et al., Mol.

Cell. Biol. 5:410-413 (1985)) by digestion with HaeII and gel electrophoresis. Synthetic adaptors are added onto this fragment to convert the HaeII ends into HindIII ends and the resulting fragment is joined to pUC19 digested with HindIII. The resulting plasmid pUCH contains the hygromycin cassette such that transcription of hph and beta-lactamase are in the opposite orientation. A 1.3 kb SalI fragment containing a DHFR gene driven by SV40 transcriptional signals is isolated from pTND (Connors et al., DNA 7:651-661 (1988)) by digestion with SalI and gel electrophoresis. This fragment is ligated to pUCH digested with SalI. The resulting plasmid pUCD contains the DHFR cassette such that DHFR and are transcribed in the same direction. A 1.76 kb BamHI fragment from the phage F15 (Friezner Degen et al., J. Biol. Chem. 261:6972-6985 (1986)) which contains 1.45 kb of DNA flanking the transcriptional start of human tissue plasminogen activator (t-PA) in addition to the first exon and part of the first intron is isolated by gel electrophoresis after BamHI digestion. This fragment is joined to pUCD following digestion of the latter with BamHI. The resulting plasmid pUCG has the promoter of the t-PA fragment oriented opposite to that of the DHFR cassette. The t-PA fragment contains a single NcoI site, which is not unique to pUCG. A partial NcoI digest is carried out and a NotI linker is inserted. The resulting plasmid pCG contains a unique NotI site in the t-PA fragment which allows the plasmid to be linearized prior to transformation of the primary human diploid fibroblasts in order to increase the frequency of homologous recombination (Kucherlapati et al., Proc. Natl. Acad. Sci. USA 81:3153-3157 (1984)).

35

Preparation of Primary Recipients

The plasmid pCG linearized with NotI is introduced into the primary recipients by

lectroporation employing DNA at 10nM. The resulting cells are then grown in selective medium (EEMEM with 200 µg/ml hygromycin B). Resistant colonies are isolated and analyzed by PCR (Kim and Smithies, 5 Nucleic Acids Res. 16:8887-8903 (1988)) using as primers the sequences GCGGCCTCGGCCTCTGCATA and CATCTCCCCTCTGGAGTGGA to distinguish homologous integrants from random ones. Amplification of cellular DNA by PCR using these two primers yields a 10 fragment of 1.9 kb only when DNA from correctly targeted cells is present. Cells comprising the DHFR gene integrated into the t-PA region are expanded and used as a source of genetic material for preparation of secondary recipients.

15

#### Preparation of Secondary Recipients

Metaphase chromosomes are prepared Nelson et al., J. Mol. Appl. Genet. 2:563-577 (1984) from recipients demonstrating homologous recombination with the DHFR and are then transformed in DHFR-deficient CHO cells by calcium phosphate mediated gene transfer (Nelson et al., J. Mol. Appl. Genet. 2:563-577 (1984)). The cells are then grown in selective medium (alpha-medium containing 200 µg/ml hygromycin B). 20 Resistant colonies are isolated and analyzed for expression of human t-PA (Kaufman et al., Mol. Cell. Biol. 5:1750-1759 (1985)). The cell clones are then grown in selective medium containing progressively higher concentrations of methotrexate (.02-80 µM, with 25 steps of 4-fold increases in concentration). After this amplification procedure, the cells are harvested and the human t-PA is analyzed employing an ELISA assay with a monoclonal antibody specific for t-PA (Weidle and Buckel, Gene 51:31-41 (1987)). Clones 30 providing for high levels of expression of t-PA are stored for subsequent use.

Isolation of a Genomic Clone Containing  
Sequences for Targeting  
Erythropoietin

A clone was obtained by screening a human  
5 placental DNA genomic library (Clontech) in EMBL 3-  
SP6/T7 using two 36 bp oligonucleotide probes 5'-  
CTGGGTTGCTGAGTTCCGCAAAGTAGCTGGGTCTGG-3' and 5'-  
CGGGGGTCGGGCTGTTATCTGCATGTGTGCGTGCG-3' to the  
presumed promoter region of human erythropoietin.  
10 From this clone two subclones were created in pSP72  
(Krieg and Melton (1987) Meth. Enzymol. 155, 397-415),  
one containing a 5 kb BamHI-HindIII fragment from the  
region upstream to the coding region of EPO (pTD.1)  
and one containing a 5 kb HindIII-BamHI fragment  
15 coding for EPO (pTD.2).

Construction of DNA Fragment for  
Targeting Erythropoietin

A plasmid pCG.1 was constructed by  
20 replacement of the polylinker of pBluescript SK(-)  
(Stratagene) between the SacI and KpnI sites with a  
synthetic double stranded 72 base pair DNA fragment  
(FIG. 1). Referring to FIG. 2, into pCG.1 was cloned  
between the HindIII and XbaI sites a 678 bp fragment  
25 containing the enhancer and promoter of the immediate  
early gene of human cytomegalovirus (CMV, Boshart et  
al (1985) Cell 41, 521-530) obtained by a PCR  
amplification of the plasmid pUCH.CMV (gift of  
M. Calos, Stanford U.) using the oligonucleotide  
30 primers 5'-  
CGCCAAGCTTGGCCATTGCATACGTT-3' and 5'-  
GAGGTCTAGACGGTCACTAAACGAGCTCT-3' in order to engineer  
HindIII and XbaI sites respectively onto the ends of  
the resultant fragment. The resultant plasmid pCG.CMV  
35 was used for further constructions.

The 620 bp BstEII-XbaI fragment from the  
pTD.2 was joined by the use of a BstEII-XbaI adapter  
to pCG.CMV restricted with XbaI to create the plasmid

pCG.CMV/EPO, in which the BstEII site of the EPO fragment is next to the promoter end of the CMV fragment. Into pCG.CMV/EPO was cloned successively a 1.94 kb fragment encoding methotrexate resistance from the plasmid pSV2dhfr (Subramani *et al* (1981) Mol. Cell. Biol. 1, 854-864) and a 1.15 kb fragment encoding G418 resistance from the plasmid pMC1neo polyA (Thomas and Capecchi (1987) Cell 51, 503-512). The neo gene was obtained as an XhoI-SalI fragment and the dhfr gene was obtained by PCR amplification using the primers 5'-  
GGACGCGTGGATCCAGACATGATAAGATA-3' and 5'-  
GGACGCGTCAGCTGTGGAAATGTGTGTCAG-3' designed to add MluI sites at the ends of the resultant fragment. The neo and dhfr genes were cloned into the XhoI and MluI sites respectively of pCG.CMV/EPO to give the plasmids pCG.CMV/EPO/DHFR and pCG.CMV/EPO/Neo/DHFR such that their transcription is in the same orientation as that of CMV. Finally, the 5 kb BamHI-HindIII fragment from pTD.1 was added via ClaI adapters at the ClaI site of pCG.CMV/EPO/Neo/DHFR to give pCG.HR1. In pCG.HR1, the 5' 5kb EPO fragment is in the same orientation as that of the 620 bp BstEII-XbaI fragment with respect to the original lambda clone.

A 9.54 kb fragment containing the 5' 5kb BamHI-HindIII EPO fragment, the dhfr and G418 markers, the CMV enhancer/promoter and the 620 bp BstEII-XbaI EPO fragment can be released from pCG.HR1 as a NotI or SacII fragment. This NotI fragment can be used for homologous recombination as it is designed to serve as an omega structure in recombination having 5 kb and 620 bp of homology to facilitate the event (FIG. 3).

For electroporation, the DNA was first cut with NotI, then extracted with phenol/chloroform and precipitated by the addition of ethanol before centrifugation. The resultant DNA pellet was resuspended at a concentration of 2 mg/ml in a volume (10  $\mu$ l) of 10 mM Tris-HCl, 1 mM EDTA (TE).

Introduction of DNA into cells

Transformed primary human 293 embryonal kidney cells (ATCC CRL 1573) were cultured in Cellgro 5 DMEM H16 (Mediatech) supplemented with 10% calf serum, glutamine (2 mM) and penicillin (100 U/ml)/streptomycin (0.1 mg/ml) and grown at 37°C in 5% CO<sub>2</sub>. At 90% confluence, cells were prepared for electroporation by trypsinization, concentration by brief centrifugation and resuspension in PBS at 10<sup>7</sup> 10 cells/0.8 ml. The cells were equilibrated at 4°C, and DNA (50 µg) restricted with NotI (as described above) was added. The mixture was electroporated at 960 µF and 260 V with a BioRad Gene Pulser and then iced again for 10 min before plating onto a 10 cm dish. After incubation at 37°C for 48 hr, the cells from a 15 10 cm dish were split equally among 5 24-well plates in media containig G418 at 0.6 mg/ml (effective concentration). Under these electroporation 20 conditions, 4-10 colonies/well survive drug selection after 2 weeks.

Detection of Homologous Recombination by PCR Analysis

Using NotI restricted DNA from pCG.HR1, 25 successful homologous recombination is obtained by insertion of the 3.8 kb construct at the targeted EPO locus while simultaneously deleting 1.2 kb of genomic sequence (FIG. 3). PCR is used to detect unique targeting events versus random integration of the DNA 30 as diagrammed in FIG. 4. Two primers are synthesized, one to the 3' end of CMV and the other to the region 3' to the XbaI site used for the 620 bp BstEII-XbaI fragment in the targeting DNA. A homologous recombination event generates a DNA target in the 35 genome from which these primers produce an amplification product of 860 bp.

In order to detect the targeting event, pools of clones (from the electroporated 293 cells)

from 4 wells each (representing about 16 colonies) were generated by trypsinizing wells and using 90% of each well for the pool. The remaining 10% of each well was then reseeded back into the well. Genomic DNA was then prepared from each pool as follows. The cells in each pool were pelleted by centrifugation for 2 min. in a 1.5 ml microcentrifuge tube, resuspended in PBS (20  $\mu$ l), and treated for 1 hr at 37°C with a solution (400  $\mu$ l) containing 10 mM Tris-HCl (pH 7.5), 100 mM NaCl, 5 mM EDTA, 1% SDS and RNase A (40  $\mu$ g/ml). Proteinase K (10  $\mu$ l, 10 mg/ml) was then added, and the samples were incubated for 4 hr at 50°C before extractions by vigorous vortexing with phenol/chloroform (200  $\mu$ l each), then with chloroform (400  $\mu$ l), the addition of ethanol (800  $\mu$ l), and centrifugation at 25°C for 10 min. The DNA pellets were washed with 70% ethanol, dried and resuspended in TE (20  $\mu$ l). An average of 40  $\mu$ g of genomic DNA was obtained from each sample.

Approximately 1  $\mu$ g from each sample of genomic DNA was used for PCR analysis. The DNA in a volume (10  $\mu$ l) of TE was boiled for 10 min. prior to the addition of PCR mix (40  $\mu$ l). The reaction (50  $\mu$ l) contained 10 mM Tris-HCl (pH 9.0 at 25°C), 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.01% gelatin, 0.1% Triton X-100, 200  $\mu$ M dNTPs, 1  $\mu$ M each of the primers 5'-AAGCAGAGCTCGTTAGTGAACCG-3' and 5'-TGAGCGTGAGTTCTGTGGAATGTG-3', and 1.5 U of Taq DNA polymerase (Promega). Following an initial incubation of 94°C for 3 min, the samples were subjected to 45 cycles of denaturation at 94°C for 1 min., annealing at 66°C for 1.5 min. and extension at 72°C for 2 min. At the end of the 45 cycles, the samples were incubated an additional 5 min. at 72°C. A portion (20  $\mu$ l) of each sample was analyzed on a 1% agarose gel run in TBE and stained with ethidium bromide. Out of the 90 pools analyzed from 3 electroporations, two samples were identified which exhibited the correct

size fragment by ethidium bromide staining. The DNA from the PCR reaction was recovered and subjected to restriction mapping with XbaI. The correct amplification product should upon treatment with XbaI yield two fragments, 669bp and 191bp. The samples from the two pools both yield fragments of the correct sizes. In addition, the sample from pool 1 exhibits other bands in the uncut material.

Following the procedure described previously, metaphase chromosomes are prepared from the recipients demonstrating homologous recombination with DHFR and transformed in DHFR deficient CHO cells. After isolating resistant colonies and analyzing for expression of EPO, the cell clones are grown in selective medium containing progressively higher concentrations of methotrexate (.02-80  $\mu$ M) with steps of 4-fold increases in concentration. The cells are then harvested, cloned and screened for production of EPO. Clones providing for at least 2-fold enhancement of EPO production are isolated.

It is evident from the above results, that the subject method provides for a novel approach to expression of a wide variety of mammalian genes of interest. The method is simple, only requires the knowledge of a sequence of about 300 bp or more in the region of a target gene, and one may then use substantially conventional techniques for transferring the amplifiable region to an expression host, and production of the desired product in high yield.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill in

20

the art in light of the teachings of this invention  
that certain changes and modifications may be made  
thereto without departing from the spirit or scope of  
the appended claims.

5

WHAT IS CLAIMED IS:

1. A method for producing mammalian  
5 proteins comprising:

growing mammalian secondary expression host  
cells comprising multiple copies of an amplifiable  
region comprising a target gene heterologous to said  
secondary expression host and expressing a protein of  
10 interest and an amplifiable gene, whereby said target  
gene is expressed and said protein is produced;

wherein said secondary host expression cells  
are produced by the method comprising:

transforming primary mammalian cells  
15 comprising said target gene with a construct  
comprising an amplifiable gene and at least one  
flanking region of a total of at least about 150 bp  
homologous with a DNA sequence at the locus of the  
coding region of said target gene to provide  
20 amplification of said target gene, wherein said  
amplifiable gene is at a site which does not  
interfere with the expression of said target gene,  
whereby said construct becomes homologously integrated  
into the genome of said primary cells to define an  
25 amplifiable region;

selecting for primary cells comprising said  
construct by means of said amplifiable gene or other  
marker present in said construct;

isolating DNA portions of said genome from  
30 said primary cells, wherein said portions are large  
enough to include all of said amplifiable region;

transforming secondary expression host cells  
with said primary cell DNA portions and cloning said  
transformed secondary expression host cells to  
35 produce clones of said secondary expression host  
cells differing in said DNA portions present in said  
secondary expression host cells;

selecting clones of said mammalian secondary expression host cells comprising said amplifiable region; and

5 amplifying said amplifiable region by means of an amplifying agent, wherein said amplifying is prior to said isolating or after said selecting and prior to said growing.

2. A method according to Claim 1, wherein said amplifiable gene is a mammalian DHFR gene.

10 3. A method according to Claim 1, wherein said portions are metaphase chromosomes.

4. A method according to Claim 1, wherein said portions are restriction fragments.

15 5. A method according to Claim 1, wherein said primary cells are human cells.

6. A method according to Claim 5, wherein said human cells are fibroblast cells.

20 7. A method according to Claim 1, wherein said construct comprises a biocidal marker providing resistance to a biocide for said primary host cells.

8. A method for producing mammalian proteins comprising:

25 transforming mammalian primary mammalian cells comprising said target gene with a construct comprising an amplifiable gene and at least one flanking region of at least about 150 bp homologous with a DNA sequence within 50 kb of the coding region of said target gene, wherein said amplifiable gene is at a site which does not interfere with the 30 expression of said target gene, whereby said construct becomes homologously integrated into the genome of said primary cells to define an amplifiable region comprising said amplifiable gene and said target gene in said genome;

35 selecting for primary cells comprising said construct by means of said amplifiable gene or other marker present in said construct;

isolating DNA portions of said genome from said primary cells, wherein said portions are large enough to include all of said amplifiable region;

transforming mammalian secondary expression host cells with said primary cell DNA portions, wherein said secondary expression host cells are of a different species from said primary host cells, and cloning said transformed secondary expression host cells to produce clones of said secondary expression host cells differing in said DNA portions present in said secondary expression host cells;

selecting clones of said mammalian secondary expression host cells comprising said amplifiable region;

amplifying said amplifiable region by means of an amplifying agent, wherein said amplifying is prior to said isolating or after said selecting; and

growing said secondary expression host cells comprising multiple copies of said amplifiable region, whereby said target gene is expressed and said protein is produced.

9. A method according to Claim 8, wherein said amplifying is with said secondary expression host cells.

25 10. A method according to Claim 8, wherein said primary cells are human cells.

11. A method according to Claim 10, wherein said human cells are diploid fibroblast cells.

30 12. A method according to Claim 8, wherein said amplifiable gene is a mutated DHFR gene having a higher Km than the wild-type gene.

13. A method according to Claim 12, wherein said secondary host expression cell is DHFR deficient.

35 14. A method according to Claim 8, wherein said construct further comprises a marker gene separated from said amplifiable region by an homologous flanking region.

15. A human cell comprising an amplifiable gene at other than its wild-type site in the human genome and within the locus of a target gene expressing a protein to provide amplification of said target gene.

5 16. A human cell according to Claim 14, wherein said cell is a normal cell.

10 17. A human cell according to Claim 14, wherein said cell is a neoplastic cell.

18. A human cell according to Claim 14, wherein said amplifiable gene is a DHFR gene.

19. A mammalian cell other than a human cell for expression of mammalian proteins in culture comprising an amplifiable region comprising an amplifiable gene within 10kb of a human wild-type gene expressing a protein, wherein said two genes are separated by substantially solely human wild-type sequence associated with said target gene and the flanking sequence associated with the amplifiable gene.

20 20. A method for producing, cells for expression of a heterologous protein in culture, said method comprising:

transforming mammalian primary cells comprising said target gene with a construct comprising an amplifiable gene and at least one flanking region of at least about 150bp homologous with a DNA sequence within 10kb of the coding region of said target gene, wherein said amplifiable gene is at a site which does not interfere with the expression of said target gene, whereby said construct becomes homologously integrated into the genome of said primary cells to define an amplifiable region comprising said amplifiable gene and said target gene in said genome;

35 selecting for primary cells comprising said construct by means of said amplifiable gene or other marker present in said construct;

isolating DNA portions of said genome from said primary cells, wherein said portions are large enough to include all of said amplifiable region;

transforming mammalian secondary expression host cells with said primary cell DNA portions, wherein said secondary expression host cells are of a different species from said primary host cells, and cloning said transformed secondary expression host cells to produce clones of said secondary expression host cells differing in said DNA portions present in said secondary expression host cells;

selecting clones of said mammalian secondary expression host cells comprising said amplifiable region; and amplifying said amplifiable region by means of an amplifying agent, wherein said amplifying is either prior to said isolating or after said selecting.

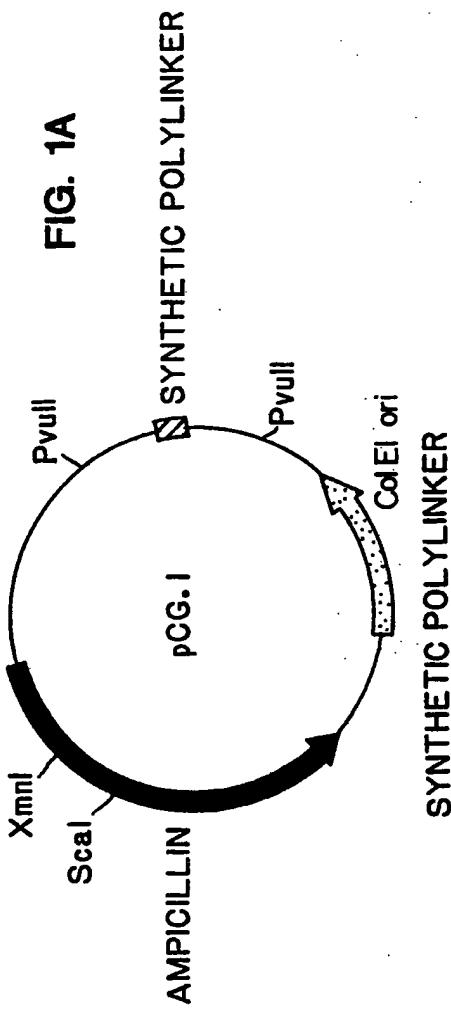
21. A method according to Claim 20, wherein said amplifying is with said secondary expression host cells.

22. A method according to Claim 20, wherein said primary cells are human cells.

23. A method according to Claim 22, wherein said human cells are diploid fibroblast cells.

24. A method according to Claim 20, wherein said amplifiable gene is a mutated DHFR gene having a higher Km than the wild-type gene.

25. A method according to Claim 24, wherein said secondary host expression cell is DHFR deficient.



**FIG. 1B**

NotI	SacII	OlaI	MluI	XbaI	HindIII	XbaI	BstEII	SacII	NotI
TGGGGCCGGTTATCGATGACGGGTCTCAGCAAGCTTCTAGACGGTCACCTTCCGGGGCGCTGTAC	TCGAACGCCGAAATAGCTACTGCCAGGCTCGTCAAGAGATCTGCCAGTGGAAAGGCA								

2 / 6

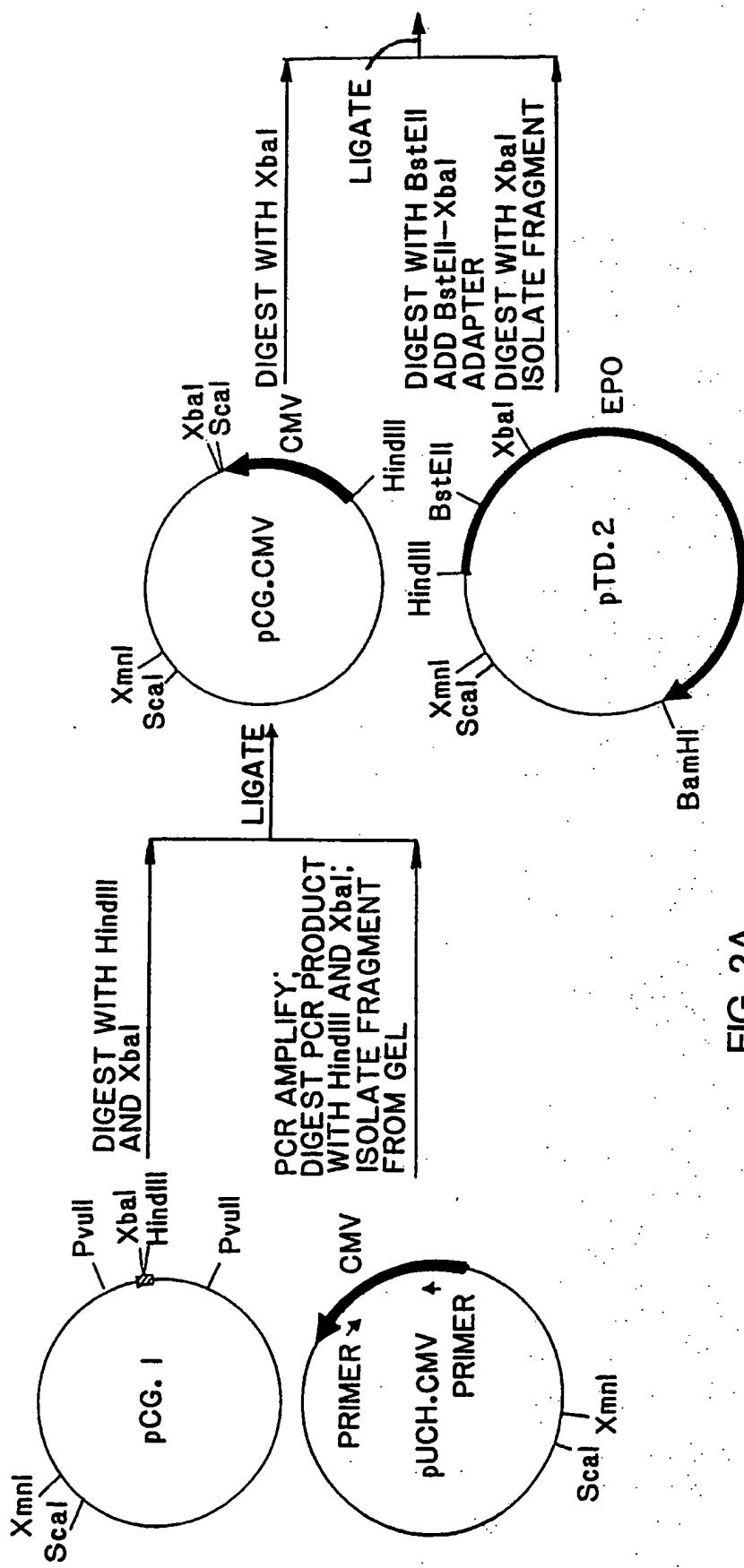


FIG. 2A

SUBSTITUTE SHEET

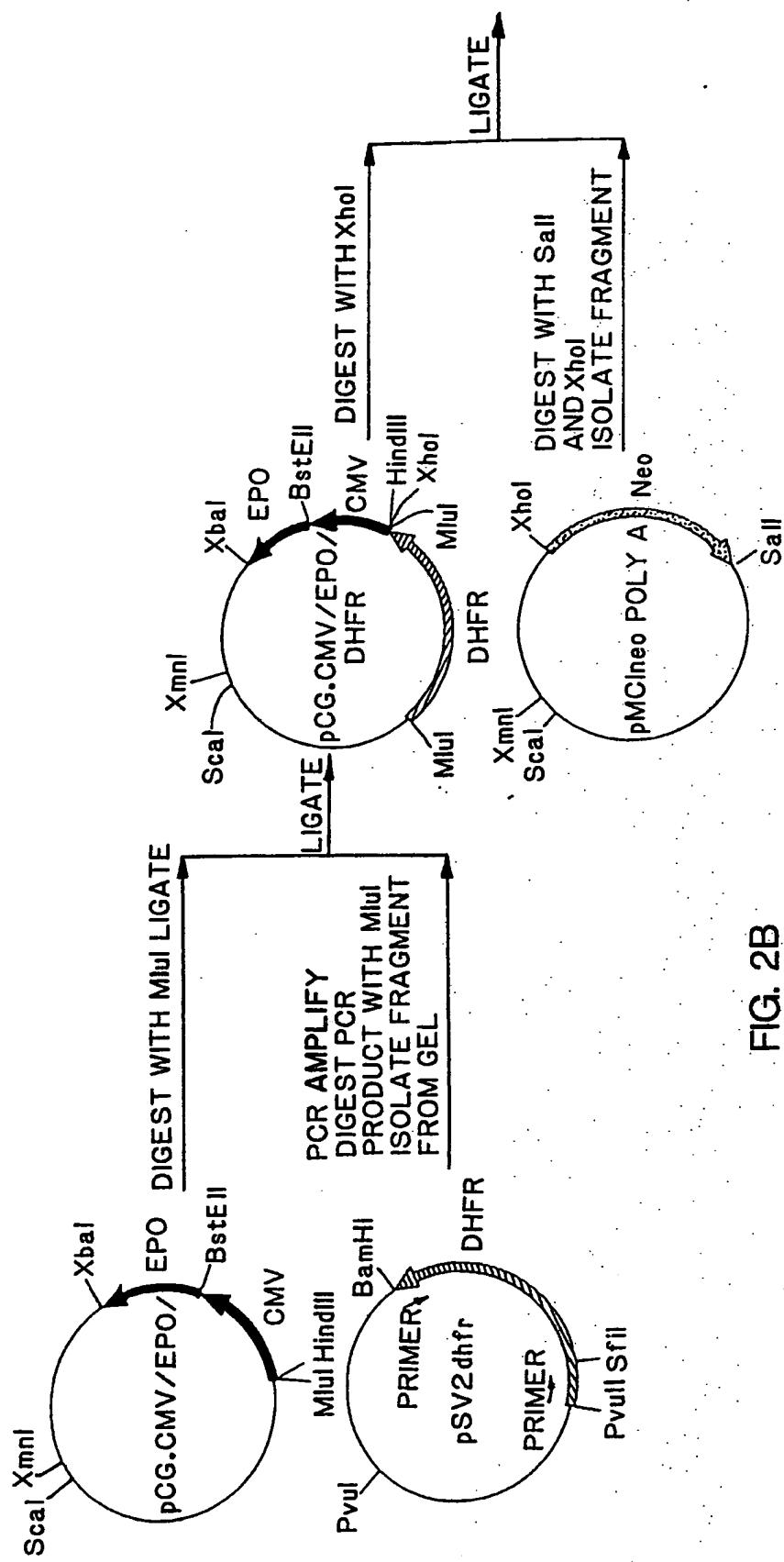


FIG. 2B

SUBSTITUTE SHEET

4 / 6

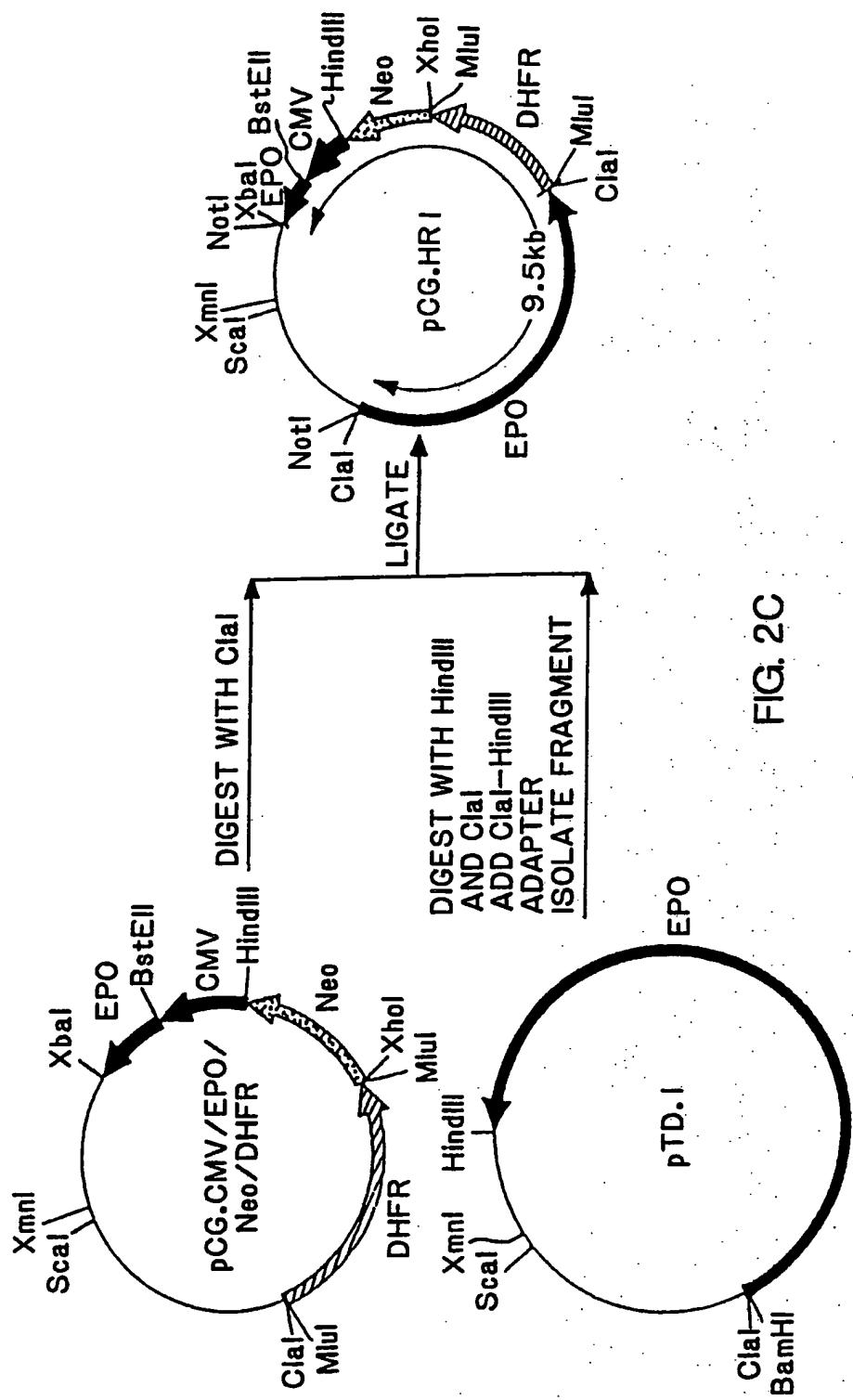


FIG. 2C

SUBSTITUTE SHEET

5 / 6

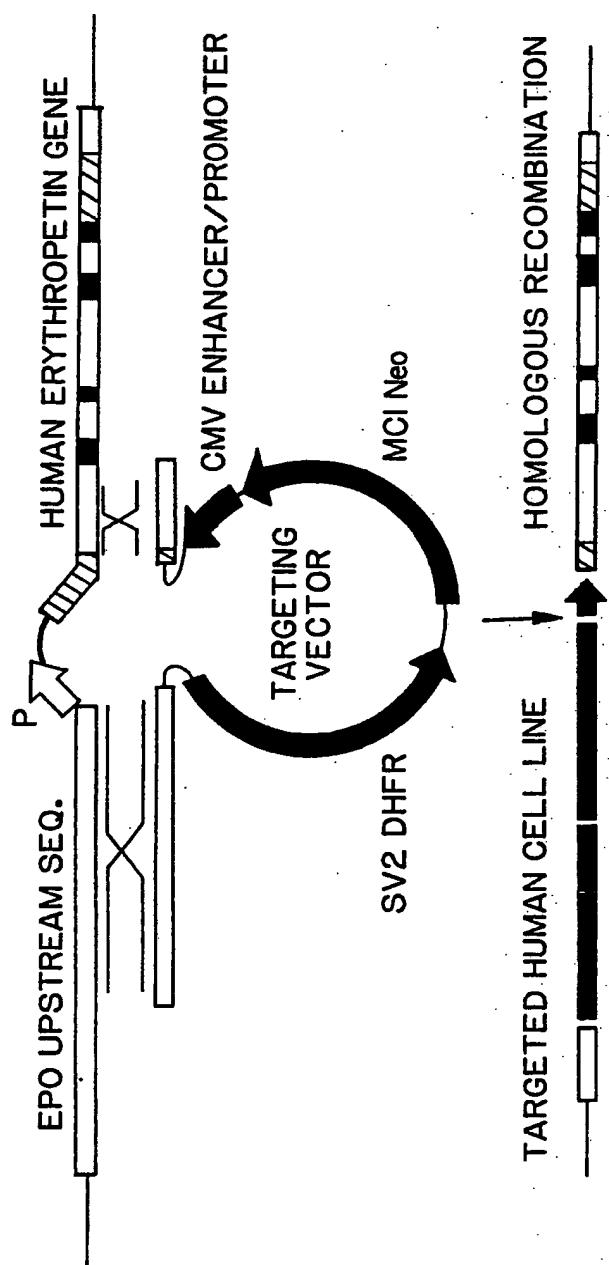


FIG. 3

SUBSTITUTE SHEET

6 / 6

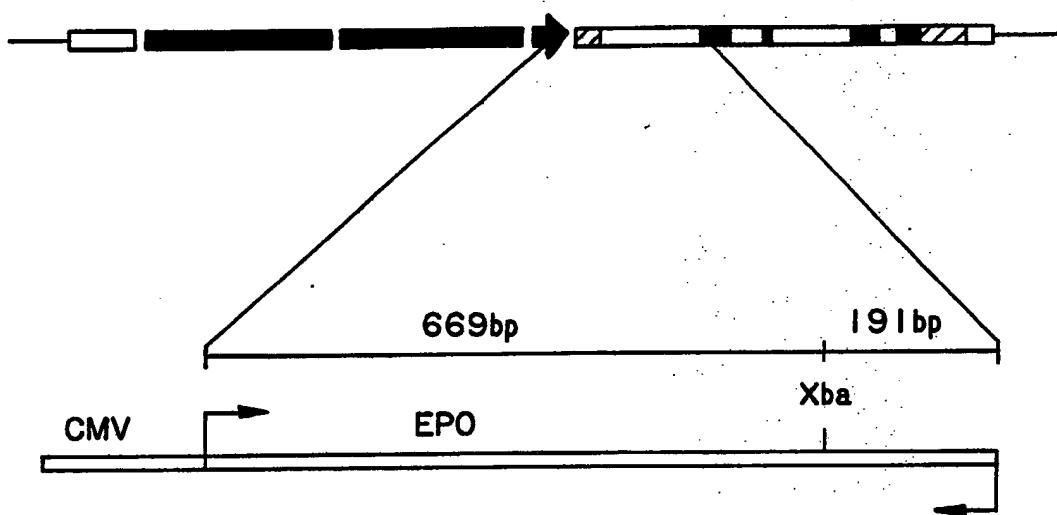


FIG. 4

# INTERNATIONAL SEARCH REPORT

International Application No PCT/US90/06425

## I. CLASSIFICATION OF SUBJECT MATTER (If several classification symbols apply, indicate all) <sup>15</sup>

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(S): C12P 21/06; C12N 15/00

US: 435/69.1,172.3,320; 935/1,6,9,10,11,12,16,24,32,34,42,60,93,109

## II. FIELDS SEARCHED

Minimum Documentation Searched <sup>16</sup>

Classification System	Classification Symbols
US	435/69.1,172.3,320; 935/1,6,9,10,11,12,16,24,32,34,42,60,93,109
Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched <sup>17</sup>	

## APS, CAS, Dialog

## III. DOCUMENTS CONSIDERED TO BE RELEVANT <sup>18</sup>

Category <sup>19</sup>	Citation of Document, <sup>16</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No. <sup>18</sup>
Y	Proc. Natl. Acad. Sci., Volume 77, issued July 1980, G. Urlaub et al., "Isolation of chinese hamster cell mutants deficient in dihydrofolate reductase activity, pages 4216-4220, see table I and results.	12,24
Y	US, A, 4,656,134 (Ringold) 07 April 1987, see entire document & col. 5, lines 37-56.	2-4,6,11, 15,16,18 & 23
Y	Proc. Natl. Acad. Sci., Vol. 84, issued October 1987, Song et al. "Accurate modification of a chromosomal plasmid by homologous recombination in human cells" pages 6820-6824, see entire article.	3

### \* Special categories of cited documents: <sup>15</sup>

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the International filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

## IV. CERTIFICATION

Date of the Actual Completion of the International Search <sup>20</sup>

09 December 1990

Date of Mailing of this International Search Report <sup>20</sup>

12 FEB 1991

International Searching Authority <sup>20</sup>

ISA/US

Signature of Authorized Officer <sup>20</sup>

Suzanne Ziska

## FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

Y	Proc. Natl. Acad. Sci., Volume 84, issued August 1987, Kamarck <u>et al.</u> , "Carcinoembryonic antigen family: Expression in a mouse L-cell transfecant & characterization of a partial cDNA in bacteriophage agt11", pages 5350-5354, see entire article.	1,3-5,7-9, 10,13,17, 19-22 & 25
---	--	---------------------------------------

V.  OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE<sup>1</sup>

This International search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1.  Claim numbers ....., because they relate to subject matter<sup>1</sup> not required to be searched by this Authority, namely:

2.  Claim numbers ....., because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out<sup>1</sup>, specifically:

3.  Claim numbers ....., because they are dependent claims not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI.  OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING<sup>2</sup>

This International Searching Authority found multiple inventions in this international application as follows:

1.  As all required additional search fees were timely paid by the applicant, this International search report covers all searchable claims of the International application.

2.  As only some of the required additional search fees were timely paid by the applicant, this International search report covers only those claims of the International application for which fees were paid, specifically claims:

3.  No required additional search fees were timely paid by the applicant. Consequently, this International search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4.  As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

## Remark on Protest

- The additional search fees were accompanied by applicant's protest.
- No protest accompanied the payment of additional search fees.

**III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SEC ND SHEET)**

Category <sup>16</sup>	Citation of Document, <sup>16</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No <sup>18</sup>
Y	Mol. Cell Biol. Volume 4, 1984, J. Rubnitz et al., "The Minimum Amount of Homology Required for Homologous recombination in Mammalian Cells", pages 2253-2257, see entire article.	1,8,14, 19 & 20